

Multiple Skalierung nach dem Kristallisationsprinzip

Eine Alternative zur explorativen Faktorenanalyse

Multiple Scaling According to the Principle of Crystallization

An Alternative Method to Exploratory Factor Analysis

Thomas Müller-Schneider

Herzog-Max-Straße 38, D-96047 Bamberg

Zusammenfassung: In diesem Beitrag geht es um eines der zentralen methodischen Probleme, die bei der sozialwissenschaftlichen Skalierung auftreten: das Erkennen mehrdimensionaler Datenstrukturen. In der Forschungspraxis setzt man dazu in aller Regel die Faktorenanalyse ein. Dieses Verfahren führt aber aufgrund seiner restriktiven Modellannahmen nicht immer zum gewünschten Ergebnis. Anhand eines Fallbeispiels zeige ich, dass komplexe Dimensionen unter Umständen nur „orthogonal fragmentiert“ wiedergegeben werden. Ein alternatives Verfahren, das ohne störende Modellrestriktionen auskommt, ist die von Mokken (1971) entwickelte multiple Skalierung dichotomer Items. Die Mokkenskalierung verfährt nach dem Kristallisationsprinzip und entdeckt zuverlässig mehrdimensionale Strukturen. Die auf dem Kristallisationsprinzip aufbauende multiple Skalierung ist ein allgemeines mehrdimensionales Skalierungsverfahren, das mit Hilfe der hierarchischen Clusteranalyse auch im Rahmen der klassischen Testtheorie eingesetzt werden kann.

1. Einleitung

In der empirischen Sozialforschung werden spezielle Verfahren eingesetzt, um Skalen zu konstruieren, die latente Dimensionen qualitativ erfassen und quantitativ messen sollen. Einstellungen, Handlungsbereitschaften, Präferenzen und Fähigkeiten sind typische Gegenstandsbereiche solcher Dimensionen. Eines der wesentlichen methodischen Probleme, die bei der sozialwissenschaftlichen Skalierung auftreten, ist das Erkennen mehrdimensionaler Strukturen. Dabei geht es um die beiden miteinander zusammenhängenden Fragen, ob ein bestimmtes inhaltliches Universum, wie beispielsweise „alltagsästhetische Vorlieben“, nur eine einzige oder mehrere inhaltlich unterscheidbare Dimensionen aufweist, und wie diese einzelnen Dimensionen gegebenenfalls anhand von erhobenen Indikatoren identifiziert werden können.

Dieser Problembereich steht im Mittelpunkt des nachfolgenden Beitrags. Zur Verdeutlichung der Thematik gehe ich zunächst auf die Eigenschaften eindimensionaler und mehrdimensionaler Datenstrukturen und im Zusammenhang damit auf Skalierungsverfahren ein. Dabei werde ich hervorheben, dass man zur Entdeckung mehrdimensionaler Strukturen methodische Verfahren einsetzen muss, die Itemgruppen auffinden können, die intern hoch und extern niedrig korrelieren (Abschnitt 2). In

der Forschungspraxis wird zu diesem Zweck meist die explorative Faktorenanalyse eingesetzt, die aber aufgrund ihrer restriktiven Modellannahmen nicht immer zum gewünschten Ergebnis führt. Anhand eines Fallbeispiels aus dem Bereich der Lebensstilforschung möchte ich zeigen, dass die Faktorenanalyse komplexe Dimensionen unter Umständen nicht vollständig, sondern nur „orthogonal fragmentiert“ wiedergibt (Abschnitt 3). Mit der von Mokken (1971) entwickelten multiplen Skalierung dichotomer Items steht ein alternatives Verfahren zur Verfügung, das auf dem Kristallisationsprinzip aufbaut und mehrdimensionale Strukturen zuverlässig auffindet. Dieses Verfahren ist eine verallgemeinerbare Skalierungsprozedur, die mit Hilfe der hierarchischen Clusteranalyse auch im Rahmen der klassischen Testtheorie eingesetzt werden kann (Abschnitt 4). Im Fazit werden die wesentlichen Ergebnisse der Untersuchung zusammengefasst (Abschnitt 5).

2. Dimensionale Datenstrukturen und Skalierungsverfahren

Ausgangspunkt für die weiteren Ausführungen ist eine eindimensionale Datenstruktur. Deren empirische Grundlage bildet eine Reihe von Items, meist standardisierte verbale Reiz- und Reaktionsvorgaben, die im Interview solche Alltagssituationen

simulieren, in denen eine theoretisch relevante Dimension sichtbar wird. Ein hochkultureller Stil beispielsweise lässt sich durch Items erfassen, die u. a. nach der Häufigkeit von Theater- und Konzertbesuchen, dem Interesse für das „Feuilleton“ und der Vorliebe für klassische Musik fragen. Bei einzelnen Personen macht sich die zugrunde liegende Dimension als konsistentes Antwortmuster bemerkbar. Wer hochkulturell orientiert ist, wird der Tendenz nach auf alle relevanten Indikatoren „positiv“ reagieren. Wer umgekehrt nichts mit Hochkultur anfangen kann, wird insgesamt eher ablehnend reagieren. Bezogen auf viele Fälle bedeutet dies, dass die gemessenen Indikatoren hoch miteinander korrelieren. In Deutungs- und Wissensgemeinschaften tritt eine latente Dimension demzufolge als abgrenzbare Gruppe statistisch eng zusammenhängender Items in Erscheinung.

Dieser korrelationsstatistische Ansatz entspricht dem in der klassischen Testtheorie verwendeten Konzept von Homogenität. Darunter versteht man allgemein die spezifische Eigenart einer Gruppe von Items, dieselbe Eigenschaft von Personen zu messen. Homogenität zeigt sich im Rahmen der klassischen Testtheorie durch hohe Interkorrelationen der betreffenden Items. Je homogener eine Skala ist, desto größer ist die statistische Maßzahl für ihre interne Konsistenz (Heidenreich 1984: 365f, 370). Eine wichtige Maßzahl der internen Konsistenz, die auf der durchschnittlichen Korrelation aller Items einer Skala basiert, ist Cronbachs Alpha (McIver/Carmines 1981: 29). Das klassische Homogenitätskonzept ist zum Teil auch in der modernen, probabilistischen Testtheorie zu finden. Zu nennen ist hierbei das auf dichotome Items ausgelegte Skalierungsverfahren nach Mokken (1971), bei dem der Koeffizient „Loevingers H“ als Kriterium für die Homogenität eines Itemsatzes herangezogen wird (Jansen 1982). Dieses Homogenitätsmaß ist ein gewichtetes Mittel aller paarweisen Itemkorrelationen H_{ij} ,¹ die mit dem korrigierten Phi-Koeffizienten für Vier-Felder-Tafeln (Φ_{kor}) identisch sind (Stokman/van Schuur 1980: 23). Eine homogene Mokkenskala weist für alle Itempaare eine deutlich positive Korrelation auf.

¹ Für zwei dichotome Items (i und j) gilt: $H_{ij} = 1 - F_{ij}/E_{ij}$. F_{ij} entspricht der Zahl der beobachteten Fehler im Sinne der Guttman-Skalierung (Zahl der Befragten, die das leichtere Item „verneinen“ und das schwierigere „bejahen“). E_{ij} ist die Zahl der erwarteten Fehler bei statistischer Unabhängigkeit der beiden Items i und j.

Loevingers H berechnet sich wie folgt:

$$H = 1 - \frac{\sum_{i=1}^{k-1} \sum_{j=i+1}^k F_{ij}}{\sum_{i=1}^{k-1} \sum_{j=i+1}^k E_{ij}}$$

Die Beschreibung der eindimensionalen Datenstruktur als abgrenzbare Gruppe hoch korrelierender Items lässt sich nun unmittelbar auf die mehrdimensionale Datenstruktur ausdehnen. Teilt sich ein gegebenes inhaltliches Universum in mehrere Dimensionen auf, so hat man es nicht mit einer, sondern mit entsprechend vielen Itemgruppen zu tun, die intern hohe und extern geringe Korrelationen aufweisen. Jeder dieser Cluster repräsentiert eine eigenständige Dimension. Daraus ergibt sich unmittelbar, wie man vorgehen muss, wenn man mehrdimensionale Strukturen erkennen will: Eine gegebene Matrix von Interitemkorrelationen ist nach Gruppen mit den genannten Eigenschaften zu durchsuchen und zu ordnen. Aus den gefundenen Gruppen können dann mit dem methodischen Instrumentarium eindimensionaler Ansätze, etwa der klassischen Testtheorie, die jeweils endgültigen Skalen zur Messung einer bestimmten Dimension entwickelt werden. Die Analyse mehrdimensionaler Datenstrukturen ist im vorliegenden Kontext somit letztlich nicht auf „echte“ mehrdimensionale, sondern auf multiple eindimensionale Konstrukte ausgerichtet. Letztgenannte haben einen hervorgehobenen methodologischen Stellenwert und sind, folgt man McIver und Carmines (1981: 14), für die sozialwissenschaftliche Theoriebildung besser geeignet: „The final and most important reason why unidimensional scaling models continue to be of substantial interest is that they are isomorphic with the primary type of concepts devised by social scientists. ... In other words, using unidimensional scaling models to measure unidimensional concepts puts the theory construction and the measurement strategy on the same analytical level.“

Bei sehr kleinen Korrelationsmatrizen lassen sich homogene Itemgruppen durch bloße Dateninspektion ausfindig machen, bei größeren Matrizen hingegen stößt diese Vorgehensweise jedoch schnell an ihre Grenzen, so dass spezielle Verfahren zur explorativen Datenanalyse eingesetzt werden müssen. Eindimensionale Skalierungsverfahren kommen dafür nicht in Frage, da sie nicht auf die Erkennung mehrdimensionaler Itemstrukturen ausgelegt sind.² In der Forschungspraxis wird hauptsächlich die Faktorenanalyse verwendet, die man auch als multidimensionale Verallgemeinerung der klassischen Testtheorie betrachten kann (Fischer 1974: 77). Bei

² Dies gilt insbesondere für die Raschskalierung, deren Homogenitätskonzept nichts mit dem oben explizierten zu tun hat, sondern als „lokale stochastische Unabhängigkeit“ definiert ist (vgl. Stelzl 1979: 653f).

ihrem explorativen Gebrauch treten allerdings Anwendungsprobleme auf, die im folgenden Abschnitt behandelt werden.

3. Zur Anwendungsproblematik der explorativen Faktorenanalyse

Unter dem Sammelbegriff „Faktorenanalyse“ werden heute viele, zum Teil sehr verschiedene multivariate Analysetechniken verstanden. Grundsätzlich lassen sich zwei Modelle der Faktorenanalyse unterscheiden, das der gemeinsamen Faktoren und die Hauptkomponentenanalyse. Beide Verfahren unterscheiden sich zwar hinsichtlich ihrer „Modellphilosophie“, führen aber in der Anwendung meist zu sehr ähnlichen Ergebnissen (Velicer/Jackson 1990). In der soziologischen Forschung dominiert die Hauptkomponentenanalyse, weshalb hier nur dieses Verfahren berücksichtigt wird. Neben ihrer konfirmatorischen Variante kann die Faktorenanalyse auch zur Exploration komplexer Korrelationsmatrizen eingesetzt werden, mit dem Ziel intern hoch korrelierende Variablencluster zu identifizieren. Die explorative Faktorenanalyse erfolgt in zwei Schritten. Im ersten wird aus einem gegebenen Satz von Items eine bestimmte Anzahl von Faktoren extrahiert, die (nach der Hauptachsenmethode) so konstruiert werden, dass sukzessive die (verbleibende) Varianz aller Items maximal aufgeklärt wird. Bei der mathematischen Extraktion werden die Faktoren allerdings nicht notwendigerweise so gewählt, dass die jeweiligen Faktorladungen die zu analysierenden Itemgruppen trennscharf wiedergeben. Um dies zu erreichen, werden die Faktoren in einem zweiten Schritt um ihr Achsenkreuz rotiert. Im Idealfall ergibt sich dadurch eine „Einfachstruktur“, bei der alle Variablen eines homogenen Clusters nur auf einem Faktor hoch laden und auf allen anderen überhaupt nicht (Gaensslen/Schubö 1976: 238ff).

Die forschungspraktisch typische Anwendung der explorativen Faktorenanalyse enthält zwei restriktive Modellvorgaben. Die erste ist „Orthogonalität“, d.h. die durch Extraktion und Rotation identifizierten Faktoren sind im statistischen Sinne wechselseitig voneinander unabhängig. Die zweite Restriktion besteht darin, dass zwischen Variablen bestehende Ähnlichkeitsstrukturen überhaupt nur dann analysiert werden können, wenn zuvor die Anzahl der zu extrahierenden Faktoren festgelegt oder mit Hilfe eines statistischen Kriteriums bestimmt wurde. Die Faktorenanalyse sucht also nach Itemclustern, die sich in einem festgelegten, n -dimensionalen

Raum orthogonaler Faktoren beschreiben lassen (vgl. Jacoby 1991: 53). Bei einer Clusterstruktur, die dieser Modellannahme weitgehend entspricht, ist die Faktorenanalyse zweifellos ein geeignetes Verfahren, um zugrunde liegende mehrdimensionale Strukturen zuverlässig zu erkennen. Die in sozialen Kollektiven vorfindbaren Dimensionen sind jedoch keine „orthogonalen“, wohlseparierten und eindeutig abgrenzbaren Sinngebilde, sondern unscharfe, zum Teil miteinander zusammenhängende und ineinander übergehende Phänomene. Auf der Datenebene resultieren daraus vielfältig überlappende Korrelationsstrukturen, die dann ihrerseits dazu führen können, dass die explorative Dimensionsanalyse unerwünschte Ergebnisse hervorbringt.

Wie aus der Methodenforschung bekannt ist, führt die orthogonale Rotation zu dimensional Artefakten, wenn die zugrunde liegenden Konstrukte stark miteinander korreliert sind (Kim/Mueller 1990: 37). Zusätzlich kann sich die Anzahl der zu extrahierenden Faktoren verzerrend auf die theoretisch relevanten Ergebnisse auswirken, selbst wenn die Annahme der Orthogonalität nicht schwerwiegend verletzt ist. Trivial ist dabei der Fall, dass „zu wenige“ Faktoren extrahiert werden. Es ist offensichtlich, dass dann nicht alle bedeutsamen Dimensionen erkannt werden können. Eine Faustregel zur Bestimmung der Faktorenanzahl gibt deshalb vor, im Zweifelsfall eher mehr als weniger Faktoren zu extrahieren und die „unnötigen“ dann bei der Weiterverarbeitung der Ergebnisse auszusondern (Gaensslen/Schubö 1976: 226). Diese Regel entspricht dem in der Forschungspraxis häufig angewandten Vorgehen, alle Faktoren zu berücksichtigen, deren Eigenwert größer als eins ist. Dadurch werden aber „zu viele“ Faktoren extrahiert, insbesondere wenn – wie das bei der explorativen Datenanalyse typischerweise der Fall ist – große Variablenzahlen analysiert werden (Bortz 1985: 662). Obwohl das „overfactoring“ als weniger gravierender Modellfehler betrachtet wird als das „underfactoring“ (Fabrigar et al. 1999: 278), sind „überzählige“ Faktoren durchaus problematisch: Je mehr Faktoren extrahiert und rotiert werden, desto eher werden einzelne Fragmente homogener Itemgruppen artifiziell als orthogonale Faktoren ausgewiesen. So entsteht die Gefahr, dass die relevanten Items einer interessierenden Dimension nicht komplett von einem einzigen Faktor erfasst, sondern gewissermaßen „orthogonal fragmentiert“ und auf verschiedene Faktoren „verteilt“ werden.

Dieses bislang zu wenig beachtete Anwendungsproblem soll an einem Fallbeispiel aus dem Bereich der Lebensstilforschung konkretisiert und verdeutlicht

werden. Ausgangspunkt sind drei von Schulze (1992) konstruierte alltagsästhetische Schemata, das Hochkulturschema, das Trivialschema und das Spannungsschema. Bei diesen drei Geschmacksstilen handelt es sich um komplexe situationsübergreifende Reaktionstendenzen, die hier als theoretisch relevante Dimensionen betrachtet werden. Für die vorliegende Untersuchung wurden sie aus einer Gesamtzahl von 64 alltagsästhetisch relevanten Items rekonstruiert, die in einer Zusatzbefragung des Wohlfahrtssurveys 1993 erhoben wurden.³ Bei der Dimensionsanalyse dieser Items kam es darauf an, möglichst homogene Itemstrukturen im Sinne der klassischen Testtheorie zu identifizieren. Die entsprechenden Maßzahlen finden sich in Tabelle 1. Wie die hohen Werte für Cronbachs Alpha zeigen, sind alle drei Gesamtskalen sehr konsistent. Die Trennschärfekoeffizienten weisen beim Trivialschema und beim Spannungsschema Werte von mindestens 0.5 auf, und auch beim Hochkulturschema überschreiten die meisten Items dieses Niveau; keiner der Trennschärfekoeffizienten ist kleiner als 0.46. Dies weist auf eine durchweg hohe Itemhomogenität innerhalb der Skalen hin. Wie die hier nicht wiedergegebenen Interskalenkorrelationen belegen (keine ist größer als 0.23), sind die drei alltagsästhetischen Schemata weitgehend voneinander unabhängig, d.h. sie lassen sich annähernd als orthogonale Dimensionen darstellen.

Dieser dreidimensionalen Datenstruktur werden nun die Ergebnisse einer am selben Datensatz durchgeführten Hauptkomponentenanalyse gegenübergestellt. Die Zahl der zu extrahierenden Faktoren wurde mit Hilfe des erwähnten und in der Lebensstilforschung häufig verwendeten Eigenwertkriteriums (z.B. Georg 1998: 147) bestimmt. Es wurden insgesamt 15 Faktoren extrahiert und anschließend nach dem üblichen Varimax-Kriterium rotiert, das nach einer orthogonalen Einfachstruktur im Faktorenraum sucht. Nicht die einzelnen statistischen Ergebnisse der Faktorenanalyse sind hier bedeutsam, sondern das Verhältnis der drei Dimensionen zu den extrahierten Faktoren wie es aus Tabelle 1 hervorgeht. Das Trivialschema wird vollständig von einem einzigen Faktor abgebildet. Alle dazugehörigen Items laden hoch auf dem ersten Faktor. Das Hochkulturschema wird noch überwie-

gend von einem, nämlich dem zweiten Faktor erfasst, wobei allerdings eines der dazugehörigen Items („Bücher lesen“) auf keinem der 15 rotierten Faktoren lädt, ein weiteres („Psychologische Literatur“) auf dem siebten und ein Fragment von drei Items („Fachbücher lesen“, „Interesse für Dokumentationen“ und „Politische Magazine“) auf dem dritten Faktor. Dieser zeichnet sich außerdem durch hohe Ladungen von Items aus, die sich auf das Interesse bei Zeitungslektüre beziehen („Wirtschaftsteil“, „Außenpolitik“ und „Innenpolitik“). Auf dem siebten Faktor laden neben der Neigung zu „psychologischer Literatur“ und zu „esoterischer Literatur“ zwei weitere Items, die auch vom zweiten Faktor erfasst werden: die Lektüre von „Gedichten“ und „Biografien“. Noch deutlicher als beim Hochkulturschema zeigt sich beim Spannungsschema eine Fragmentierung durch „zu viele“ orthogonal rotierte Faktoren. Diese Stilvariable wird gleichmäßig auf zwei scheinbar unabhängige Faktoren „aufgeteilt“, wobei sich der eine durch actionorientierten TV-Konsum (vierter Faktor) und der andere durch die Vorliebe für „junge“ Musikstile (fünfter Faktor) auszeichnet.

Wie das vorangegangene Fallbeispiel verdeutlicht, führen die Restriktionen der Faktorenanalyse unter bestimmten Umständen zu gravierenden methodischen Artefakten. Im Prinzip lassen sich diese problematischen Restriktionen auch innerhalb des faktorenanalytischen Modellrahmens umgehen. Die Orthogonalitätsannahme kann durch die oblique (schiefwinklige) Rotation von Faktoren aufgelöst werden, so dass – zumindest modelltheoretisch – auch hoch korrelierte Konstrukte bzw. Skalenfragmente abbildbar sind. Der Einfluss der Faktorenzahl lässt sich ausschalten, indem man Analysen für eine große Spannweite unterschiedlicher Faktorenzahlen berechnet und dann anhand der gefundenen Ergebnisse die „richtige“ auswählt. Diese Vorgehensweise setzt – wenn die jeweiligen Faktorenstrukturen nicht gleichzeitig oblique rotiert werden – allerdings voraus, dass die zu findenden Dimensionen einen weitgehend orthogonalen Raum aufspannen⁴. Die angeführten Modellverallgemeinerungen bringen eine enorme und unpraktikable Verkomplizierung der explorativen Faktorenanalyse mit sich. Schon allein die oblique Faktorenrotation wird in der soziologi-

³ Die mehrstufigen Items umfassen Freizeittätigkeiten, den Musikgeschmack, das Interesse für Fernsehsendungen, die Vorliebe für Literaturgattungen sowie das Interesse für Zeitungssparten (vgl. Spellerberg 1993). In die Analyse gingen nur Befragte aus den alten Bundesländern (N = 1550) ein.

⁴ Im Fallbeispiel lassen sich die drei alltagsästhetischen Schemata mit einer entsprechend vorgegebenen Zahl orthogonal rotierter Faktoren auffinden. An diesem Punkt wird besonders deutlich, dass man unter gewissen Bedingungen in die Faktorenanalyse „hineinstecken“ muss, was man eigentlich von ihr erwartet: Informationen über die Dimensionalität eines Itempools.

schen Forschungspraxis kaum angewendet. Hinzu kommt, dass etliche Verfahren zur schiefwinkligen Rotation zur Auswahl stehen, mit zusätzlichen Modellparametern, die vielfältige Wahlmöglichkeiten und somit auch willkürliche Entscheidungen zulassen (vgl. Tschopp 1991: 51). Inhaltlich-theoretische Ergebnisse können erheblich von der jeweils gewählten Methode und Modellspezifikation beeinflusst werden. Eine wesentlich einfachere und gleichzeitig zuverlässigere Vorgehensweise wird im nächsten Abschnitt dargelegt.

4. Multiple Skalierung nach dem Kristallisationsprinzip

Das nachfolgend vorgestellte Verfahren zur Entdeckung von Clustern intern hoch und extern niedrig korrelierender Items stammt von Robert J. Mokken (1971). Es basiert auf einer Skalierungsprozedur, die nach dem Kristallisationsprinzip funktioniert und zur Konstruktion möglichst homogener Skalen dichotomer Items entwickelt wurde. Bei dieser Prozedur geht man schrittweise nach einem Algorithmus vor, der sich in groben Zügen wie folgt beschreiben lässt:

1. Suche das homogenste Itempaar eines gegebenen Pools (bemessen am Korrelationskoeffizienten H_{ij} bzw. $\Phi_{i, \text{kor}}$; siehe Abschnitt 2). Betrachte die beiden Items als potentiellen Kristallisationskern einer Skala.
2. Wähle aus dem restlichen Pool von Items dasjenige aus, das die höchste Skalierbarkeit für das entstehende Skalenfragment erbringt (in Bezug auf Loevingers H). Erweitere den Kristallisationskern um dieses Item.
3. Wiederhole den unter 2. genannten Schritt so lange, bis ein bestimmter Schwellenwert von H erreicht ist, der nicht unterschritten werden soll. Die agglomerative Konstruktion der Skala ist dann abgeschlossen.

Diese drei Verfahrensschritte sind dazu geeignet, eine eindimensionale Datenstruktur abzubilden. Um auch mehrdimensionale Strukturen erfassen zu können, geht Mokken (1971: 194f) den Weg der „multiplen Skalierung“. Darunter ist eine Skalierungsprozedur zu verstehen, in deren Verlauf eindimensionale Skalen gebildet werden, und zwar eine nach der anderen, indem das Kristallisationsprinzip zur Skalenbildung immer wieder auf den jeweils noch verbliebenen Pool von Items angewendet wird. Auf diese Weise lassen sich die in einem Datensatz vorhandenen Itemcluster sukzessive

identifizieren. Zur Durchführung einer multiplen Skalierung ist der obige Algorithmus um einen entscheidenden iterativen Schritt erweitern:

4. Versuche, aus den restlichen Items eine weitere Skala zu bilden. Beginne dazu wieder bei Schritt 1.

Die Mokkenskalierung ist – wenn es um die Entdeckung mehrdimensionaler Strukturen geht – der Faktorenanalyse deutlich überlegen. Denn das multipl wiederholte Kristallisationsprinzip verlangt weder, dass schon vor der eigentlichen Analyse eine bestimmte Anzahl aufzufindender Dimensionen festgelegt wird, noch enthält es irgendwelche modellbedingten Vorgaben über das Verhältnis der Dimensionen untereinander. Daraus ergeben sich im Vergleich zur explorativen Faktorenanalyse zwei grundlegende Vorteile. Erstens maximiert der von potentiellen Kristallisationskernen ausgehende Skalierungsprozess die Chance, große Felder homogener Items und damit theoretisch relevanter Dimensionen nicht nur als Fragmente, sondern in ihrer gesamten Extension zu erfassen. Zweitens können die gefundenen Skalen auf unmittelbare Weise so miteinander korrelieren, wie es den tatsächlichen Verhältnissen entspricht. Eine komplizierte und für Fehlspezifikationen anfällige oblique Rotation von Faktoren ist nicht vonnöten.

Um eine Mokkenskalierung der bereits faktorenanalytisch untersuchten 64 Items vornehmen zu können, mussten diese zunächst dichotomisiert werden.⁵ Anschließend wurde die beschriebene Skalierungsprozedur mit dem Programm „MSP“ (Debets/Brouwer 1988) durchgeführt, wobei der Schwellenwert für den Skalierbarkeitskoeffizienten H mit .45 so gewählt wurde, dass er nach den von Mokken (1971: 185) festgelegten Grenzen deutlich über dem Mindestwert für eine Skala „mittlerer“ Kristallisation liegt.⁶ Die dimensionsanalytischen Kennzahlen der drei endgültigen Mokkenskalen, die den bereits vorgestellten alltagsästhetischen Schemata entsprechen (vgl. Tabelle 1), finden sich in Tabelle 2. Überaus homogen ist der Itemsatz, der das Trivialschema repräsentiert, wohingegen die Skalierbarkeitskoeffizienten für das Hochkultur- und Spannungsschema etwas niedriger sind. Neben Loevingers H sind weiterhin die Werte für den Ko-

⁵ Gewählt wurde eine „inhaltliche“ Dichotomisierung, die sich an Bedeutungen der Antwortkategorien orientiert und diese im Sinne einer „positiven“ bzw. „negativen“ Antwort interpretiert. Dadurch wurden die unterschiedlichen Itemschwierigkeiten beibehalten. Die Mittelkategorie „teils/teils“ wurde als positive Antwort gewertet.

⁶ $0,5 \leq H$: starke Skala; $0,4 \leq H < 0,5$: mittlere Skala; $0,3 \leq H < 0,4$: schwache Skala.

Tabelle 1 Alltagsästhetische Schemata: Dimensionsanalytische Kennzahlen nach der klassischen Testtheorie

| Dimension | Items (Vorliebe, Interesse für...) | Trennschärfe | Faktor(-ladungen) | | |
|-----------------------|---------------------------------------|--------------|-------------------|------|------|
| Trivialschema | | | FA 1 | | |
| | Volksmusik | .74 | .79 | | |
| | Schlager | .70 | .73 | | |
| | Blasmusik | .64 | .74 | | |
| | TV: Shows/Quiz | .53 | .56 | | |
| | TV: Volkstheater | .68 | .77 | | |
| | TV: Heimatfilme | .73 | .79 | | |
| Cronbachs Alpha = .87 | | | | | |
| Hochkulturschema | | | FA 2 | FA 3 | FA 7 |
| | Bücher lesen | .63 | | | |
| | Theater/Konzerte | .46 | .51 | | |
| | Klassische Musik | .48 | .77 | | |
| | Oper | .66 | .80 | | |
| | Moderne Literatur | .62 | .47 | | |
| | Klassische Literatur | .70 | .62 | | |
| | Zeitung: Kulturteil | .57 | .54 | | |
| | TV: Kultursendungen | .53 | .56 | .33 | |
| | TV: Politische Magazine | .56 | | .71 | |
| | TV: Dokumentationen | .47 | | .63 | |
| | Fachbücher | .50 | | .45 | .38 |
| | Psychologische Literatur | .47 | | | .74 |
| | Gedichte | .50 | .47 | | .50 |
| | Biographien | .60 | .43 | | .44 |
| Cronbachs Alpha = .88 | | | | | |
| Spannungsschema | | | FA 4 | FA5 | |
| | TV: Actionfilme | .58 | .76 | | |
| | TV: Science Fiction | .67 | .84 | | |
| | TV: Horrorfilme | .54 | .69 | | |
| | Science Fiction, Fantasy-Romane | .56 | .62 | | |
| | Comics | .50 | .33 | .51 | |
| | Popmusik | .56 | | .56 | |
| | Rockmusik | .62 | | .76 | |
| | Punk, Heavy | .50 | | .69 | |
| | TV: Pop- u. Rockmusik | .63 | | .69 | |
| Cronbachs Alpha = .85 | | | | | |

Quelle: Wohlfahrtssurvey 1993 (N = 1550). Zur besseren Übersichtlichkeit wurden nur Faktorladungen der insgesamt 15 extrahierten Faktoren wiedergegeben, die größer als 0.3 sind.

Tabelle 2 Alltagsästhetische Schemata: Dimensionsanalytische Kennzahlen der Mokkenskalisierung

| Dimension | Items (Vorliebe, Interesse für...) | H_i | Faktor ¹ |
|--------------------|---------------------------------------|-------|---------------------|
| Trivialschema | Volksmusik | .59 | FA 1 |
| | Schlager | .61 | FA 1 |
| | Blasmusik | .54 | FA 1 |
| | TV: Shows/Quiz | .53 | FA 1 |
| | TV: Volkstheater | .54 | FA 1 |
| | TV: Heimatfilme | .59 | FA 1 |
| | Arzt-/Schicksalsbücher | .46 | FA 1 |
| | Zeitung: Lokalnachrichten | .46 | FA 11 |
| Loevingers H = .55 | | | |
| Hochkulturschema | Theater/Konzerte | .42 | FA 2 |
| | Klassische Musik | .47 | FA 2 |
| | Oper | .47 | FA 2 |
| | Klassische Literatur | .47 | FA 2 |
| | Zeitung: Kulturteil | .41 | FA 2 |
| | TV: Kultursendungen | .46 | FA 2 |
| | TV: Politische Magazine | .45 | FA 3 |
| | TV: Dokumentationen | .49 | FA 3 |
| | Zeitung: Innenpolitik | .54 | FA 3 |
| | Zeitung: Außenpolitik | .56 | FA 3 |
| | Biographien | .40 | FA 7 |
| Loevingers H = .46 | | | |
| Spannungsschema | TV: Actionfilme | .46 | FA 4 |
| | TV: Science Fiction | .52 | FA 4 |
| | TV: Horrorfilme | .47 | FA 4 |
| | Science Fiction, Fantasy-Romane | .46 | FA 4 |
| | Popmusik | .56 | FA 5 |
| | Rockmusik | .45 | FA 5 |
| | Punk, Heavy | .50 | FA 5 |
| | TV: Pop- u. Rockmusik | .46 | FA 5 |
| Loevingers H = .48 | | | |

1) Bei den Items, die auf mehr als einem Faktor hoch laden (vgl. Tab. 1), wurden die Faktoren mit der höchsten Ladung angegeben.
Quelle: Wohlfahrtssurvey 1993 (N = 1550).

effizienten H_i angeführt. Diese Maßzahl ist ein gewichtetes Mittel aller Korrelationen H_{ij} eines bestimmten Items (i) mit allen anderen Items⁷ und

$$H_i = 1 - \sum_{j \neq i}^k F_{ij} / \sum_{j \neq i}^k E_{ij}. \text{ Für } F_{ij} \text{ und } E_{ij} \text{ siehe Fußnote 1.}$$

kann in Analogie zur klassischen Testtheorie als Trennschärfekoeffizient interpretiert werden (Stokman/van Schuur 1980: 23). Die nach den Maßstäben der Mokkenskalisierung durchgängig hohen Werte für H_i bestätigen die bereits im Rahmen der klassischen Testtheorie festgestellte hohe Itemhomogenität innerhalb der Skalen (vgl. Abschnitt 3).

Betrachtet man die Faktoren, denen die skalierten Items in der 15-Faktoren-Lösung (vgl. Abschnitt 3) zugeordnet sind, so wird deutlich, dass die scheinbar orthogonalen Faktorenfragmente durch das Kristallisationsprinzip der Mokkenskalisierung jeweils zu einer übergreifenden, theoretisch relevanten Dimension zusammengefügt werden. Das Spannungsschema, dies zeigte sich schon bei der klassischen Testtheorie, setzt sich aus Items des vierten und fünften Faktors zusammen. Bei der Trivialekultur kommt mit der Präferenz für Lokalnachrichten sogar noch ein Item eines weiteren Faktors (11) hinzu. Das Hochkulturschema umfasst Items der drei Faktoren (2, 3 und 7), die auch bei der Skalierung nach der klassischen Testtheorie gefunden wurden; allerdings wurde nur ein Item skaliert, das auf Faktor 7 lädt. Dies liegt jedoch nicht am Prinzip der agglomerativen Skalenbildung, sondern an der mangelnden Skalierbarkeit der betreffenden Items (in Bezug auf H bzw. H_i). Auch hinsichtlich der Interskalenkorrelationen kann das von der klassischen Testtheorie erbrachte Ergebnis einer weitgehend orthogonalen Struktur der Dimensionen bestätigt werden.

Die von Mokken entwickelte Skalierungsprozedur ist nicht auf die von ihm verwendeten Maße beschränkt, sondern lässt sich mit beliebigen Zusammenhangsmaßen durchführen, beispielsweise mit dem Rangkorrelationskoeffizienten oder mit dem Produkt-Moment-Korrelationskoeffizienten. Die auf dem Kristallisationsprinzip aufbauende multiple Skalierung ist eine allgemeingültige Prozedur zur Entdeckung mehrdimensionaler Itemstrukturen, die auf dem korrelationsstatistischen Homogenitätskonzept (siehe Abschnitt 2) basieren. Wenn in einem gegebenen Itempool Gruppen intern hoch und extern niedrig korrelierender Items existieren, dann werden sie durch dieses Verfahren auch aufgefunden.

Wichtige Konsistenzmaße der in der Forschungspraxis überwiegend angewendeten klassischen Testtheorie bauen auf Produkt-Moment-Korrelationen auf, die ihrerseits wiederum auf der Basis multikategorialer Items berechnet werden. Da für diese Datenstruktur bislang kein Verfahren existiert, das explizit den Algorithmus zur multiplen Skalierung nach dem Kristallisationsprinzip anwenden würde, bietet es sich – wie oben durchgeführt – an, die betreffenden Items zu dichotomisieren und anschließend einer Mokkenskalisierung zu unterziehen. Mit einer Dichotomisierung sind allerdings Informationsverluste verbunden, die sich unter Umständen negativ auf die Reliabilität der Gesamtskala auswirken können (Debets/Brouwer 1988: 7). Je nach-

dem, welche der ursprünglichen Kategorien zusammengefasst werden, können weitere unerwünschte Effekte bei der Skalenbildung auftreten.⁸ Auch wenn nicht davon auszugehen ist, dass dadurch inhaltlich-theoretische Ergebnisse erheblich beeinflusst werden (vgl. Müller-Schneider 1993), kann man anstelle der bislang wenig verbreiteten Mokkenskalisierung ein multivariates Verfahren einsetzen, das in der Lage ist Produkt-Moment-Korrelationen zu verarbeiten und sich im Sinne einer multiplen Skalierung nach dem Kristallisationsprinzip verwenden lässt. Es handelt sich dabei um die hierarchische Clusteranalyse, die in üblichen Programmpaketen (z. B. SPSS) installiert ist.

Ganz allgemein sind Clusteranalysen Verfahren, deren primäres Ziel es ist, Objekte (hier: Variablen) in solche Gruppen einzuteilen, die intern möglichst homogen und extern möglichst heterogen sind (Bortz 1985: 685). Im Unterschied zu partitionierenden Verfahren, die eine vorab festgelegte Objektgruppierung iterativ optimieren, werden bei der hierarchisch-agglomerativen Clusteranalyse Objekte nach ihrer Ähnlichkeit schrittweise fusioniert, und zwar von der feinsten (jedes Objekt bildet eine Gruppe) bis hin zur größten Partition (alle Objekte befinden sich in einer Gruppe). Eines der Ähnlichkeitsmaße, das zur Fusionierung von Variablen herangezogen werden kann, ist der hier relevante Produkt-Moment-Korrelationskoeffizient. Für den eigentlichen Fusionierungsprozess ist weiterhin ein Verfahren festzulegen, wie die einzelnen Variablen schrittweise miteinander verschmolzen werden sollen. Zur Wahl steht eine Reihe von Algorithmen, unter denen ein bestimmtes Mittelwertverfahren, das „Average-Linkage Within Groups“-Verfahren (vgl. Bacher 1994: 270ff), für den gegebenen Untersuchungszweck am besten geeignet ist. Danach wird im ersten Fusionierungsschritt zunächst das Itempaar zusammengefasst, das den höchsten Korrelationskoeffizienten aufweist. Diesen beiden Items wird dann das Item zugeordnet, das in der neu zu bildenden Dreiergruppe die höchste durchschnittliche Korrelation ergibt – es sei denn, es findet sich ein neues Itempaar mit einer noch größeren

⁸ Im vorliegenden Fall wurden gewisse Dichotomisierungseffekte festgestellt. Wird die Mittelkategorie „teils/teils“ entgegen der obigen Vorgehensweise (vgl. Fußnote 3) als negative Antwort interpretiert, nimmt die Homogenität der Skalen ab und es sind insgesamt weniger Items skalierbar. Unabhängig von der gewählten Dichotomisierung zeigt sich allerdings eine weitgehend orthogonale Struktur der Dimensionen; die beiden Skalenvarianten der jeweiligen Dimensionen sind hoch korreliert (r liegt zwischen 0,63 und 0,66).

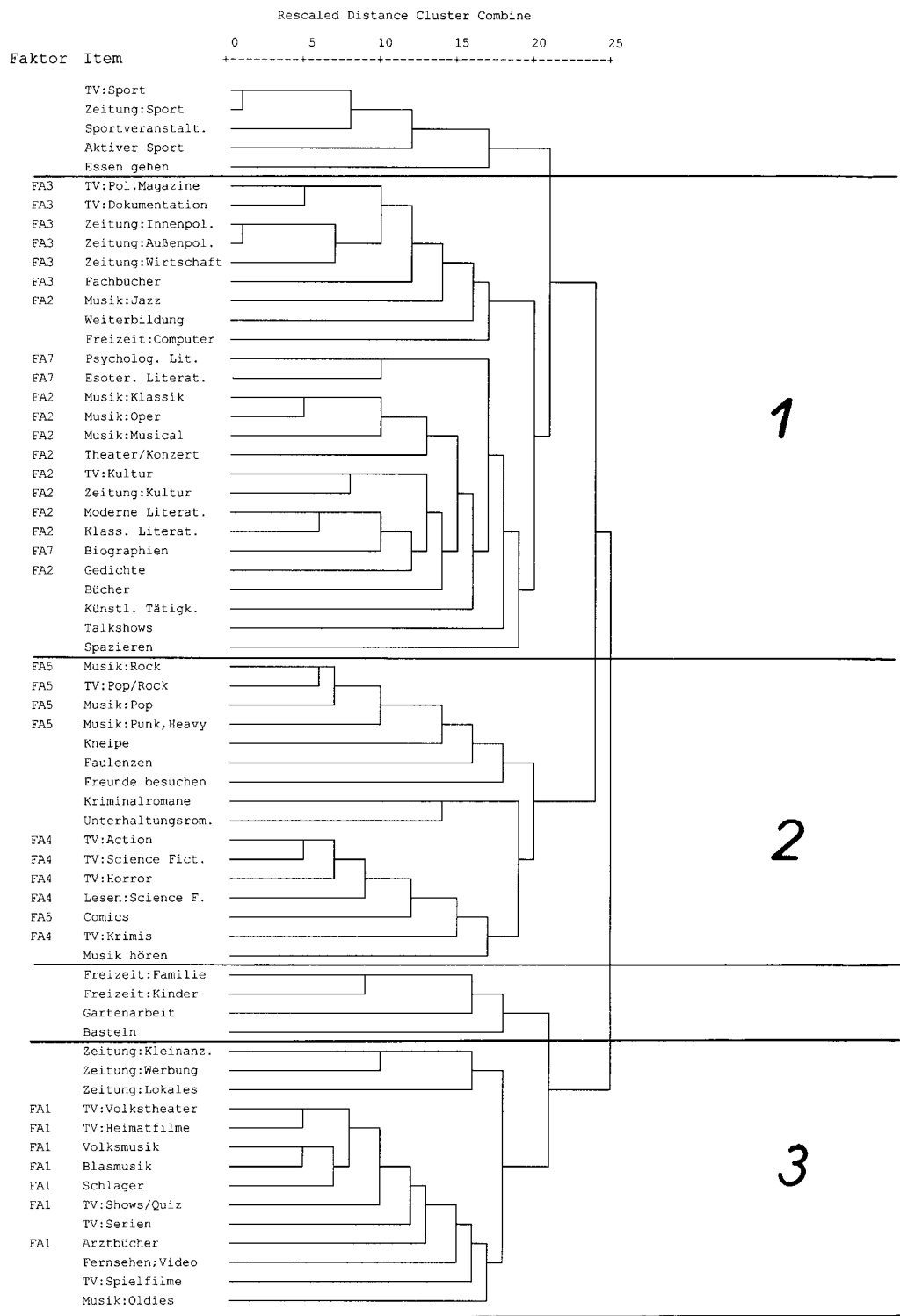


Abb. 1 Dendrogram nach dem „Average-Linkage Within Groups“-Verfahren.

Korrelation. Im Laufe der Fusionierung werden entweder weitere Itempaare gebildet oder diese um andere Items bzw. bereits gebildete Itemgruppen erweitert, und zwar immer so, dass die durchschnittliche Korrelation innerhalb der neuen Gruppe am höchsten ist.⁹ An diesem Punkt wird die große Ähnlichkeit der hierarchischen Clusteranalyse mit der auf dem Kristallisationsprinzip aufbauenden multiplen Skalierung sichtbar. Der wesentliche Unterschied liegt darin, dass der Skalierungsprozess nicht bei einem bestimmten Schwellenwert abgebrochen, sondern so lange weitergeführt wird, bis alle Variablen(gruppen) zu einer Gesamtgruppe fusioniert sind. Dennoch lassen sich wichtige Erkenntnisse zur Dimensionalität der analysierten Items gewinnen.

In Abbildung 1 wird das Dendrogramm für die mit dem „Average-Linkage Within Groups“-Verfahren durchgeführte hierarchische Klassifikation der 64 Items wiedergegeben, die bereits faktorenanalytisch untersucht wurden. Das Dendrogramm zeigt, wie die einzelnen Items, beginnend von der feinsten bis hin zur größten Partition, miteinander verschmolzen werden. In der Kopfzeile sind die durchschnittlichen Korrelationen bei den einzelnen Fusionschritten abgetragen, normiert auf die Werte zwischen 0 und 25. Je höher das im Dendrogramm ersichtliche Verschmelzungsniveau ist, desto heterogener sind die zusammengefassten Itemgruppen. So ist beispielsweise die Korrelation der beiden Items „Innenpolitik“ und „Außenpolitik“ (Interessen bei der Zeitungslektüre) deutlich höher als die durchschnittliche Korrelation der Itemgruppe, die zusätzlich das Item „Wirtschaft“ enthält. Diese Dreiergruppe wiederum ist erheblich homogener als die nächsthöhere Fusion mit den Items „Politische Magazine“ und „Dokumentationen“ (Interesse für Fernsehsendungen).

Unbeeinflusst von den erörterten Modellrestriktionen der Faktorenanalyse (Orthogonalität; festgelegte Zahl der Faktoren) lassen sich nicht nur homogene Itemgruppen erkennen, sondern auch die Ähnlichkeitsstruktur, die zwischen den einzelnen Gruppen besteht. In Abbildung 1 sind in der linken Spalte diejenigen orthogonal rotierten Faktoren

(vgl. Abschnitt 3) gekennzeichnet, die zu einem der drei Gesamtstile gehören. Die Ähnlichkeit bzw. Unähnlichkeit dieser Faktoren wird durch ihre Position in den Segmenten des Dendrogramms sichtbar. Die entscheidende Information hierfür ist das Verschmelzungsniveau. Die beiden Segmente beispielsweise, die zwischen den Faktoren 4 und 5 differenzieren, werden auf relativ niedrigem Niveau fusioniert, d.h. die genannten Faktoren sind sich hinsichtlich ihrer Interitemkorrelationen sehr ähnlich. Durchforstet man das Dendrogramm nach weiteren Ähnlichkeitsstrukturen, erkennt man, dass sich die Faktoren 2, 3 und 7 ebenfalls ähnlich sind. Die gefundenen Ergebnisse legen es nahe, die in Abbildung 1 innerhalb der drei markierten Bereiche liegenden Items jeweils insgesamt einer eindimensionalen Itemanalyse zu unterziehen. Führt man eine entsprechende Itemanalyse mit Hilfe des Trennschärfekoeffizienten durch, so erhält man die in Abschnitt 3 vorgestellten drei alltagsästhetischen Skalen, die den gefundenen Mokkenskalen sehr ähnlich sind (vgl. oben).

5. Fazit

Im Mittelpunkt der vorangegangenen Analyse stand ein wichtiges methodisches Problem, das bei der sozialwissenschaftlichen Dimensionsanalyse auftritt, nämlich das Erkennen mehrdimensionaler Itemstrukturen. Zu diesem Zweck wird hauptsächlich die explorative Faktorenanalyse eingesetzt. Dieses Verfahren ist aufgrund seiner restriktiven Modellvorgaben, insbesondere der festzulegenden Faktorenzahl und der forschungspraktisch fast durchweg unterstellten Orthogonalität, allerdings nur bedingt zur Analyse mehrdimensionaler Strukturen geeignet. Wie ich anhand eines Fallbeispiels aus dem Bereich der Lebensstilforschung gezeigt habe, besteht bei der Extraktion „zu vieler“ Faktoren das Risiko, dass theoretisch relevante Dimensionen „fragmentiert“ und auf mehrere, scheinbar voneinander unabhängige Faktoren „verteilt“ werden. In solchen Fällen wird das hier zugrunde gelegte Erkenntnisziel, übergreifende Dimensionen in einem inhaltlichen Universum zu erkennen, nicht erreicht.

Ein alternatives Verfahren, das der explorativen Faktorenanalyse deutlich überlegen ist, ist die von Mokken entwickelte multiple Skalierung nach dem Kristallisationsprinzip. Die Mokkenskalierung ist ein datenorientiertes Verfahren, das zuverlässig mehrdimensionale Strukturen entdeckt. Es basiert auf einem einfachen Algorithmus der agglomerati-

⁹ Der dargestellte Fusionierungsprozess beruht somit auf einem der klassischen Testtheorie entsprechenden Homogenitätskonzept. Auch dort ist die durchschnittliche Interitemkorrelation von ausschlaggebender Bedeutung. Besonders deutlich ist dies bei der Maßzahl Cronbachs Alpha zu sehen, in die, wie bereits erwähnt (vgl. Abschnitt 2), die durchschnittliche Korrelation der jeweiligen Skalenitems eingeht.

ven Gruppierung dichotomer Items, der den in sozialwissenschaftlich relevanten Populationen typischerweise auftretenden, vielfältig überlappenden Korrelationsstrukturen besonders angemessen ist. Das Verfahren kommt ohne unnötige und störende Modellkomplikationen aus. Weder muss die Zahl der zu findenden Dimensionen vorab festgelegt werden noch müssen diese auf irgend eine Weise „rotiert“ werden. Theoretisch relevante Dimensionen werden in ihrer Gesamtheit erfasst und können auf „natürliche Weise“ miteinander korrelieren.

Die auf dem Kristallisationsprinzip aufbauende multiple Skalierung lässt sich mit beliebigen Konsistenzkoeffizienten durchführen und ist daher ein allgemeines mehrdimensionales Skalierungsverfahren. Im Rahmen der klassischen Testtheorie kann die hierarchische Clusteranalyse eingesetzt werden, die sich im Sinne einer agglomerativen Skalenbildung anwenden lässt. Das Dendrogramm veranschaulicht ähnliche Itemgruppen bzw. Faktoren, die zusammengefasst jeweils einer eindimensionalen Itemanalyse unterzogen werden können. Im verwendeten Fallbeispiel zeigte sich, dass die Mokken-skalierung und die hierarchische Clusteranalyse im wesentlichen zum gleichen Ergebnis führen.

Literatur

- Bacher, J., 1994: Clusteranalyse. Anwendungsorientierte Einführung. München/Wien: Oldenbourg.
- Bortz, J., 1985: Statistik für Sozialwissenschaftler. Berlin/Heidelberg/New York: Springer.
- Debets, P. / Brouwer, E., 1988: Program MSP version 1.40. Amsterdam: Disc.
- Fabrigar, L. R. / Wegener, D. T. / MacCallum, R. C. / Strahan, E. J., 1999: Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods* 4: 272–299.
- Fischer, G. H., 1974: Einführung in die Theorie psychologischer Tests. Grundlagen und Anwendungen. Bern/Stuttgart/Wien: Huber.
- Gaensslen, H. / Schubö, W., 1976: Einfache und komplexe statistische Analyse. Eine Darstellung der multivariaten Verfahren für Mediziner und Sozialwissenschaftler. München/Basel: Ernst Reinhardt.
- Georg, W., 1998: Soziale Lage und Lebensstil. Eine Typologie. Opladen: Leske und Budrich.
- Heidenreich, K., 1984: Grundbegriffe der Meß- und Testtheorie. S. 352–384 in: E. Roth (Hrsg.), Sozialwissenschaftliche Methoden. Lehr- und Handbuch für Forschung und Praxis. München/Wien: Oldenbourg.
- Jacoby, W. G., 1991: Data Theory and Dimensional Analysis. Newbury Park/London/New Delhi: Sage.
- Jansen, P. G. W., 1982: Homogenitätsmessung mit Hilfe des Koeffizienten H von Loevinger: Eine kritische Diskussion. *Psychologische Beiträge* 24: 96–105.
- Kim, J.-O. / Mueller, C. W., 1990: Factor Analysis. Statistical Methods and Practical Issues. Newbury Park/London/New Delhi: Sage.
- McIver, J. P. / Carmines, E. G., 1981: Unidimensional Scaling. Beverly Hills /London: Sage.
- Mokken, R. J., 1971: A Theory and Procedure of Scale Analysis. With Applications in Political Research. The Hague/New York: Mouton.
- Müller-Schneider, T., 1993: Verschiedene Verfahren – verschiedene Ergebnisse? Vergleich der Skalierungsverfahren nach Rasch und Mokken sowie der klassischen Testtheorie am Beispiel alltagsästhetischer Schemata. *Zeitschrift für Soziologie* 22: 371–384.
- Schulze, G., 1992: Die Erlebnisgesellschaft. Kultursoziologie der Gegenwart. Frankfurt/(M.): Campus.
- Spellerberg, A., 1993: Lebensstile im Wohlfahrtssurvey 1993. Dokumentation zum Konzept und zur Entwicklung des Fragebogens. Arbeitsgruppe Sozialberichterstattung. Papers. Berlin: WZB.
- Stelzl, I., 1979: Ist der Modelltest des Rasch-Modells geeignet, Homogenitätshypothesen zu überprüfen? Ein Bericht über Simulationsstudien mit inhomogenen Daten. *Zeitschrift für experimentelle und angewandte Psychologie* 26: 652–672.
- Stokman, F. / van Schuur, W., 1980: Basic Scaling. *Quality and Quantity* 14: 5–30.
- Tschopp, A., 1990: Modellhaftes Denken in der Soziologie. Eine Untersuchung zur Anwendung formaler Modelle in der empirischen Sozialforschung und in der soziologischen Theoriebildung. Bern/Frankfurt(M.)/New York/Paris: Lang.
- Velicer, W. F. / Jackson, D. N., 1990: Component Analysis versus Common Factor Analysis: Some further Observations. *Multivariate Behavioral Research* 25: 97–114.

Summary: This paper focuses on one of the central methodological problems of scaling in the social sciences: the identification of multidimensional structures. In most practical research contexts exploratory factor analysis is used to solve this problem. However, the problem with this statistical procedure is that – due to certain restrictions – it does not always produce the desired outcome. Drawing evidence from one example, the author shows that under certain conditions complex dimensions are “fragmented” by orthogonal exploratory factor analysis. An alternative method which does not have distorting restrictions is a multiple scaling procedure for dichotomous items developed by Mokken (1971). Mokken scaling operates according to the principle of crystallization and reliably identifies multidimensional structures. Multiple scaling is a general multidimensional scaling method that, with the help of cluster analysis, can be implemented within classical test theory.